and New England, is obviously associated with the winter cyclones which appear in the southwest and move northeastward. Rain and warm weather occur in their path and snow and cold weather to the northwest. At Dubuque, Iowa, it was found [1] that precipitation in winter occurs more frequently with falling temperature than with rising. In the Plateau and Pacific States the well-known relation between precipitation and the latitudinal position of cyclones as they approach the coast is evident, especially in the marked contrast between Oregon and California. Northern LOWS are attended by warm and wet weather in Oregon and warm and dry in California; southern LOWS by cool and dry weather in Oregon and cool and wet in California. These statements are, of course, incomplete and partial and serve only to illustrate the relations suggested by the chart. It is beyond the scope of this note to enter into a discussion of the conditions under which winter precipitation occurs in the various States and sections of the country. The sole object has been to compile and present the facts of record, expressed in State averages, showing the relationship between winter temperature and precipitation departures.

TABLE 1.—*Number of times winter temperature and precipitation departures were of like and unlike signs. Only those winters were counted in which the average temperature departure for the three months, December, January, and February, was ±2° F. or more*

States	Departures of—		Percentage having like signs
	Like signs	Unlike signs	
North Atlantic:			
New England	15	4	79
New York	15	5	75
Pennsylvania	14	7	67
New Jersey	12	6	67
Maryland and Delaware	8	8	50
Sums	64	30	68
South Atlantic:			
Virginia	9	10	47
North Carolina	8	11	42
South Carolina	8	11	42
Georgia	7	15	32
Florida	7	10	41
Sums	39	57	41

[1] T. A. Blair, Local Forecast Studies—Winter Precipitation, M. W. R., 52: 79-85.

TABLE 1.—*Number of times winter temperature and precipitation departures were of like and unlike signs. Only those winters were counted in which the average temperature departure for the three months, December, January, and February, was ±2° F. or more*—Continued

States	Departures of—		Percentage having like signs
	Like signs	Unlike signs	
Lake region, Ohio Valley, and eastern Mississippi Valley:			
Michigan	14	5	74
West Virginia	11	10	52
Ohio	15	12	56
Indiana	13	12	52
Kentucky	14	6	70
Wisconsin	12	8	60
Illinois	19	17	53
Tennessee	16	9	64
Alabama	12	12	50
Mississippi	13	8	62
Sums	139	99	58
West Gulf:			
Louisiana	9	9	50
Texas	6	11	35
Sums	15	20	43
Central Plains and middle Mississippi Valley:			
Missouri	18	8	69
Kansas	13	11	54
Oklahoma	9	8	53
Arkansas	9	6	60
Sums	49	33	60
Western upper Mississippi Valley, Missouri Valley, and Rocky Mountain:			
Iowa	18	24	43
Minnesota	11	14	44
North Dakota	5	17	23
South Dakota	11	14	44
Nebraksa	12	23	34
New Mexico	10	8	56
Colorado	6	12	33
Wyoming	1	11	8
Montana	4	16	20
Idaho	9	8	53
Sums	87	147	37
South Plateau and south Pacific:			
Nevada	6	13	32
Utah	5	12	29
Arizona	7	10	41
California	1	5	17
Sums	19	40	32
North Pacific:			
Oregon	17	2	89
Washington	13	7	65
Sums	30	9	77

INTERPOLATION OF RAINFALL DATA BY THE METHOD OF CORRELATION

ERIC R. MILLER

[Weather Bureau, Madison, Wis.]

The object of this paper is to apply to a climatological problem a method already well established in other sciences. Suppose that it is wished to interpolate from observations at near-by stations the monthly rainfall at a station where observations have been taken previously. I shall use the symbol Y to refer to rainfalls at the first, X, at the others, y and x to refer to deviations from the mean rainfalls.

Think of a "scatter diagram" each point of which represents the simultaneous rainfalls, X measured on a horizontal scale, Y on a vertical scale. The "regression line" of the statistician (**6**, p. 120) has the property that the sum of the squares of the distances of the dots of the scatter diagram measured parallel to the Y axis from the regression line, is less than from any other line. Hence, under a least-squares criterion of approximation, the regression line is the "best" representation of the relation between Y and X for all amounts of rain. The following remarks will be restricted to straight regression lines, but the fitting of curved regression lines is also practiced.

The formation of the regression equation, representing algebraically the regression line, involves calculation of the standard deviations of the observed X's and Y's, and their coefficient of correlation. Concise examples of this are given in books on statistics (**8**, p. 178–179), (**6**, p. 123) and the calculation is easily carried out with the aid of Crelle's Tables (**1**).

Horton (**3**) has given some correlation coefficients calculating from 12 months taken at random. In order to ascertain the effect of change of season upon the correlation coefficient, I have calculated it for the 32 years (1897–1928) rainfall at Waupaca and Pine River, Wis. (14 miles apart), for January, when practically all rain falls in "general" storms for May, the wettest month, with many heavy thunderstorms, and for August, a month characterized by very local rain and drought.

The results were:

	January	May	August
Correlation coefficient	0.94±0.014	0.92±0.018	0.89±0.025
Average rainfall:			
Waupaca	1.15	4.06	3.65
Pine River	1.08	4.06	3.40
Standard deviation:			
Waupaca	.82	2.08	1.66
Pine River	.70	2.48	2.03

From these statistics the regression equations, expressing rainfalls at Waupaca in terms of rainfall at Pine River, are:

In deviation from the mean:

January $y=1.10\ x$
May $y=\ .77\ x$
August $y=\ .73\ x$

In total rainfall:

January $Y=1.10\ X-\ .03$
May $Y=\ .77\ X-\ .93$
August $Y=\ .73\ X-1.17$

The effect of increasing distance between stations upon the correlation coefficient is shown by the following table of correlation with May rainfall at Waupaca:

	Pine River	Grand River Locks	Portage	Beloit
	X_1	X_2	X_3	X_4
Distance from Waupaca, miles	14	40	58	126
Correlation coefficient	0.92±0.018	0.76±0.05	0.73±0.55	0.40±0.10
Mean rainfall	4.06	4.24	3.93	3.63
Standard deviation	2.48	2.20	1.87	1.79

The regression equations, expressing Waupaca rainfall in terms of the rainfalls at each of these four stations, are:

(The variables are the deviations in inches from the mean)

Distance

14 miles $y=0.77\ x_1$
40 miles $y=\ .72\ x_2$
58 miles $y=\ .81\ x_3$
126 miles $y=\ .46\ x_4$

(The variables are the monthly rainfalls in inches)

14 miles $Y=0.77\ X_1+\ .93$
40 miles $Y=\ .72\ X_2+1.01$
58 miles $Y=\ .81\ X_3+\ .88$
126 miles $Y=\ .46\ X_4+2.37$

The decrease of the correlation coefficient per mile amounts to 0.005 or 0.006.

Calculation of regression equations for two or more control stations is more complicated, but numerical examples that can be followed by any novice are given in the books referred to at the end of this article (**6**, p. 145), (**4**, p. 205), (**2**, p. 136–138). The labor is greatly reduced by the use of Miner's Tables (**5**) or Kelley's Alignment Chart (**4**, back cover) and Chiò's method of evaluating determinants (**7**, p. 71.)

As examples of such regression equations I have calculated three involving the three control stations, Pine River (14 miles south of Waupaca), New London (18 miles east), and Stevens Point (26 miles northwest.) The following table contains the statistics on which these calculations were based:

Correlation coefficients of stations in heading, with stations at left.—Monthly rainfalls for May, **32** *years,* **1897–1928**

	Waupaca	Pine River	New London	Stevens Point
	Y	X^1	X^2	X^3
Pine River	0.92±0.018			
New London	.85± .033	0.89±0.025		
Stevens Point	.92± .018	.88± .027	0.88±0.027	
Mean rainfall	4.06	4.06	4.08	3.80
Standard deviation	2.08	2.48	2.10	1.95

Regression equations—

(1) In deviation from the means:

$$y=+0.57\ x_1-0.10\ x_2+0.50\ x_3$$
$$y=+\ .80\ x_1+\ .14\ x_2$$
$$y=+\ .52\ x_1+\ .46\ x_3$$

(2) In monthly rainfalls, inches:

$$Y=+0.48\ X_1-0.10\ X_2+0.48\ X_3+.70$$
$$Y=+\ .68\ X_1+\ .14\ X_2+\ .75$$
$$Y=+\ .44\ X_1+\ .49\ X_3+\ .42$$

It will be noted that the smaller correlation between New London and Waupaca than between New London and the other controls, although the latter are farther away, has a marked effect in diminishing the New London coefficient in these multiple regression equations.

The labor of these calculations of multiple regression equations does not increase in proportion, when a number of equations are derived for different stations, based on the same controls, because the same intermediate coefficients are used again and again in the different relations.

In closing, I wish to acknowledge the cheerful assistance of Junior Observer Alfred L. Lorenz, who calculated all of the total correlations for me.

LITERATURE CITED

(1) Crelle, A. L. Rechentafeln, Berlin, 1923.

(2) Fisher, R. A. Statistical methods for research workers, 3d ed., Edinburgh, 1930.

(3) Horton, R. E. Rainfall interpolation. Mo. Wea. Rev., vol. 5, 1923, pp. 291–304.

(4) Kelley, T. L. Statistical method. New York, 1923.

(5) Miner, J. R. Tables of $\sqrt{1-r^2}$ and $1-r^2$ for use in partial correlation and in trigonometry. Baltimore, 1922.

(6) Rietz, H. L., ed. in chief. Handbook of mathematical statistics. Boston, 1924.

(7) Whittaker, E. T., & Robinson, G. The calculus of observations. London, 1924.

(8) Yule, G. U. An introduction to the theory of statistics. London, 1912.